Molecular Characterization of *Juglans regia* L. Cultivars with SSR Markers

I. Foroni[1], R. Rao[1] and K. Woeste[2]
[1]Dipartimento di Scienze del Suolo della Pianta e dell'Ambiente, Università di Napoli "Federico II", Portici, Italy
[2]USDA Forest Service, North Central Research Station, Purdue University, West Lafayette, IN, USA

Keywords: walnut, diversity, *Juglans nigra*, cultivar identification, genetic relationships

Abstract

DNA microsatellites are abundant, uniformly distributed, hypervariable, co-dominant and highly reproducible. They are a powerful and informative method to study genetic relationships and genotype identity. We used nine microsatellites (WGA4, WGA33, WGA80, WGA147, WGA148, WGA204, WGA221, WGA256, WGA275) originally isolated in black walnut (*Juglans nigra* L.) to characterize six Persian walnut cultivars; four from Europe ('Malizia', 'Blegiana', 'Parisienne' and 'Franquette') and two from the U.S. ('Serr' and 'Hartley'). The microsatellites amplified a total of 43 putative alleles. The number of alleles per locus ranged from four to eight, with a mean of 5.3, and a molecular size range from 130 bp to 296 bp. The eight microsatellite loci were sufficiently polymorphic that unique genotypes for each cultivar were easily discernable. There are many potential applications of microsatellites to walnut genetics, including pedigree analysis, breeding, population genetics, germplasm management, pollen flow analysis, and cultivar identification.

INTRODUCTION

Simple sequences repeats (SSR) are now commonly used in evolutional and conservation studies as well as in breeding programs (Brown et al., 1996). SSRs are tandemly repeated DNA sequences of two to six nucleotides. The most common repeat cores are $(CA)_n$ and $(TG)_n$ and they were also the first ones abundantly found in human genomes (Liu et al., 1995). Microsatellites are dispersed throughout the entire genome, codominant, and show a high level of polymorphism. Many studies observed the presence of SSRs in plants (Zhao and Kochert, 1992; Koebner, 2001) and other eukaryotic genomes (Hamada et al., 1982), demonstrating their informativeness and usefulness. In the last decade, because of their advantages, many authors have used SSR markers to understand the genetic variability and genetic relationships within crops such as tomato (He et al., 2003), pear (Yamamoto et al., 2002), coconut (Meerow et al., 2003), and peach (Dirlewanger et al., 2002). They allow genetic analysis in any season. For these reasons, it is now common to use SSRs together with other characters for cultivar identification.

Microsatellites are particularly useful for studying the genetic of populations (Warburton and Hoisington, 2001), knowing and understanding the genetic structure of a populations is useful for developing an optimal strategy for in situ conservation (Gomez et al., 2004), to inform breeding programs (Portis et al., 2004), for Denomination of Protected Origin certification (Gemas et al., 2004), or to better understand how the landraces adapt to different ecological and environmental stresses (Farid et al., 2000).

In the present study, our purpose was to use the power and informativeness of microsatellites method to study the genetic relationships among and genotype identity of six cultivars of *Juglans regia*.

MATERIALS AND METHODS

Plant Material

In this work we describe six *Juglans regia* cultivars belonging to the walnut germplasm collection of the Istituto Sperimentale per la Frutticoltura (ISF-CE) (Table 1).

Four genotypes were walnut cultivars from Europe ['Blegiana' (B), 'Malizia'(M), 'Parisienne' (P), 'Franquette'(F)], and two were from the US ['Hartley'(H), 'Serr'(SR)].

DNA Extraction

Young leaves were ground in liquid nitrogen, the DNA extracted and stored at -80°C. The extraction protocol of Doyle and Doyle (1987) was modified as follows: 1 g of frozen, ground leaves was added to 10 ml of preheated (50°C) 2x CTAB buffer (2% CTAB, 50mM DTT, 0.3% β-mercaptoethanol, 1.4M NaCl, 100mM Tris, 20mM EDTA, pH=8.0) and incubated at 65°C for 30 min. The aqueous solution was extracted with 20 ml (24:1) chloroform - octanol, centrifuged 15 min at 13000 rpm, and the aqueous layer retained. Two volumes of 100% cold ethanol were added to precipitate the nucleic acids. The precipitate was spooled, washed with 0.2M ammonium acetate in 75% ethanol and air-dried for 5 min. The pellet was then resuspended in 500 µl of 65°C preheated H_2O and treated with 125 µl (10 mg/ml) RNase at 37°C for 30 min. The DNA was precipitated, washed, dried, resuspended in 250 µl H_2O and quantified in a 1% agarose gel stained with ethidium bromide (10 mg/ml) against DNA lambda standards.

SSR Primers and Amplification

As an initial screen, primers designed from the sequences of clones from an enriched (GA/CT)n library of *J. nigra* partly described previously (Woeste et al., 2002) were used to amplify genomic DNA of *J. regia* cultivars. Amplification products from these reactions were separated on 2% agarose gel (not shown). Nine primers combinations that produced clear and polymorphic products without artifactual bands (WGA4, WGA33, WGA80, WGA147, WGA148, WGA204, WGA221, WGA256, WGA275) were selected to amplify the genomic DNA of the Persian walnut cultivars.

PCR reactions contained 3.3mM $MgCl_2$, 0.10mM dNTPs, 0.5 unit Taq polymerase (New England Biolab, Beverly, MA), 0.4µM unlabelled primer (reverse), 0.4µM forward primer labelled with Hex, Ned or 6 Fam consistent with matrix D on an ABI 377 automated sequencer (Applied Biosystems, Foster City, CA), 0.4% BSA, and 15ng DNA template in a final volume of 10 µl. Thermal cycling conditions were as follows: denaturation 5 min at 95°C; 35 cycles of 1 min at 94°C, 1 min at 55°C and 2 min at 72°C; and final extension 10 min at 72°C. Each amplification included a negative control reaction without DNA template. To prepare samples for loading, every amplification was diluted up to 1:4 in water and then 1.0 µl PCR product was mixed with 0.5 µl blue dextran, 2.5 ml formamide, and 0.5 µl Tamra 500-labelled molecular size standard. The PCR products were combined into two gel sets. Products within a gel set could be multiloaded to run in a single lane using a CAL 96 paper comb (The Gel Company, San Francisco, CA) (Table 2). The samples were denatured at 94°C for 1 min before loading onto an ABI 377 sequencer (Applied Biosystems, Foster City, CA) and electrophoresed in the 6% polyacrylamide gels under denaturing conditions [7 M urea, 1 X TBE buffer (90 mM Tris borate, pH 8.3, and 2mM EDTA)]. DNA fragment length were analyzed with the software GENESCAN (Applied Biosystems, Foster City, CA). Three replicate experiments were carried out for each SSR primer pair-genotype combination.

Evaluation of Polymorphisms and Data Analysis

The SSR allele composition of each sample was determined and the CONVERT (Glaubitz, 2003), MICROSAT (Minch et al., 1997) and PHYLIP (Felsenstein, 1989) programs were used to analyse the data and to construct a neighbor-joining dendrogram based on allele frequency data calculated by Nei's genetic distance (Nei, 1972).

RESULTS AND DISCUSSION

All loci detected by a single primer pair were polymorphic among the samples studied. Each sample had just one or two alleles at each locus. A total of 45 alleles were detected with a size range of 130 – 296 bp. Twentysix alleles were private to single individuals. The number of alleles per locus varied from two to eight, with a mean of 5.0.

(Table 2). The highest frequency allele (57.0%) was the 200 bp allele of WGA256, while the alleles with lowest frequency (8.0%) were 132 bp and 186 bp amplified by WGA80. The most polymorphic locus was WGA221 with eight alleles. Table 3 represents the allelic distribution observed at nine SSR loci and among the six cultivars.

The power of the loci to uniquely each genotype varied; the number of unique patterns per locus ranged from one to six, with an average of 3.1. WGA221 produced a unique allele combination for all six cultivars, while WGA80 produced only two patterns, one for 'Blegiana' and another for the other five cultivars (Table 3). The number of private alleles per sample varied from two ('Parisienne') to seven ('Blegiana'), with a mean of 2.8, thus 'Blegiana' could be distinguished from other cultivars at six loci (WGA4, WGA33, WGA80, WGA204, WGA221, and WGA275), but 'Parisienne' in three loci just (WGA33, WGA221, and WGA256) (Table 3). A unique SSR genotype was generated for all samples. The Nei's genetic distance among samples was calculated on the basis of allele frequency data, and the cluster analysis was generated using the neighbor-joining algorithm (Fig. 1). The structure of the tree was tested by 1,000 bootstraps. The unrooted tree organized all the cultivars in well separated branches.

Variety identification was, historically, based on phenotypic analyses, but in the last ten-years the need to develop DNA markers to monitor the genetic diversity among cultivars has been considered an important goal. In fact the molecular characterization of cultivars offers a quick and reproducible assay not influenced by environmental conditions.

SSRs are easy to use, codominant, and they require only a small amount of DNA. The analysis is robust and transferable. They permit a high throughput, being easily automated using capillary electrophoresis systems with fluorescent detection, and they are more and more used in plant genomic studies (Liu et al., 1996). Several reports have shown that SSRs are a good way to evaluate the genetic relationships among cultivars (Hamza et al., 2004) due to their high level of polymorphism (He et al., 2003). In fact, the average number of alleles per locus in this study was 5.0, much higher than 1.3 and 3.9 detected in *Juglans regia* with RAPDs (Nicese et al., 1998) and ISSRs (Potter et al., 2002), respectively. It is worth noting that the high number of private alleles detected using the microsatellites in this study (twenty-six) allowed the determination of a unique SSR genotype for all the analyzed cultivars. Even more interesting is the observation that a single SSR locus uniquely identifies all the cultivars included in this study:

The phylogenetic analysis highlighted the genetic distance among the cultivars, placing each very far from the others. These results confirm the power of microsatellites to diversify distinguish *Juglans regia* cultivars using SSR primers isolated in *Juglans nigra.*

The microsatellites described here in combination with other published microsatellites (Foroni et al., 2005) provide a powerful tool for genetic analysis of *J. regia.*

Microsatellites can be used not only for variety identification but also for developing an optimal strategy for in situ conservation (Gomez et al., 2004); better understanding how the landraces adapt to different ecological and environmental stresses (Farid et al., 2000); verifying paternity in breeding programs and hybrid rootstocks (Portis et al., 2004); and Denomination of Protected Origin certification (Gemas et al., 2004).

Literature Cited

Brown, S.M., Hopkins, M.S., Mitchell, S.E., Senior, M.L., Wang T.Y., Duncan R.R., Gonzales-Candelas, F. and Kresovich, S. 1996. Multiple methods for the identification of polymorphic simple sequence repeat (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. Theor. Appl. Genet. 93:190-198.

Dirlewanger, E., Cosson, P., Tavaud, M., Aranzana, M.J., Poizat, C., Zanetto, A., Arus, P. and Laigret, F. 2002. Development of microsatellite markers in peach (*Prunus percica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). Theor. Appl. Genet. 105:127-138.

Doyle, J.J. and Doyle, J.L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bul. 19:11-15.

Farid, A., O'Reilly, E., Dollard, C. and Kelsey, C.R. Jr. 2000. Genetic analysis of ten sheep breeds using microsatellite markers. Can. J. Anim. Sci. 80 (1): 9-17.

Felsenstein J. 1989. PHYLIP – Phylogeny Inference Package (Version 3.2). Cladistics 5:164–166.

Foroni, I, Rao, R., Woeste, K. and Gallitelli, M. 2005. Characterization of *Juglans regia* L. through SSR markers and evaluation of genetic relationships among cultivars and the 'Sorrento' landrace. J. Hortic. Sci. Biotech. (In press).

Gemas, V.J.V., Almadanim, M.C., Tenreiro, R., Martins, A. and Fevereiro, P. 2004. Genetic diversity in the olive tree (*Olea europapea* L. Subsp. *Europapea*) cultivated in Portugal revealed by RAPD and ISSR markers. Genet. Resour. Crop Evol. 51:501-511.

Glaubitz, J.C. 2003. Convert: (version 1.2): A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. http://www.agriculture.purdue.edu/fnr/htlm/faculty/rhodes/student%20and%20Staff/glaubitz/software.htm

Gomez, O.J., Blair, M.W., Frankow-Lindberg, B.E. and Gullberg, U. 2004. Molecular and phenotypic diversity of common bean landraces from Nicaragua. Crop Sci. 44:1412-1418.

Hamada, H., Petrino, M.G. and Kakunaga, T. 1982. A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc. Natl. Acad. Sci. USA. 79:6465-6469.

Hamza, S., Hamida, W.B., Rebai, A. and Harrabi, M. 2004. SSR-based genetic diversity among Tunisian winter barley and relationship with morphological traits. Euphytica 135:107-118.

He, C., Poysa, V. and Yu, K. 2003. Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars. Theor. Appl. Genet. 106:363-373.

Koenber, M.D.R. 2001. Contribution of DNA molecular marker technologies to the genetics and breeding of wheat and barley. p.181- 219. In: Jules Janick (ed.), Plant Breeding Reviews, vol. 21. John Wiley & sons, New York.

Liu, Z., Biyashev, R.M. and Saghai, M.M.A. 1996. Development of simple sequences repeat DNA markers and their integration into a barley linkage map. Theor. Appl. Genet. 93:869-876.

Liu, Z.W., Jarret, R.L., Kresovich, S. and Duncan, R.R. 1995. Characterization and analysis of simple sequences repeat (SSR) loci in seashore paspalum (*Paspalum vaginatum* Swartz). Theor. Appl. Genet. 91:47-52.

Meerow, A.W., Wisser, R.J., Brown, J.S., Kuhn, D.N, Schnell, R.J. and Broschat, T.K. 2003. Analysis of genetic diversity and population structure within Florida coconut (*Cocos nucifera* L.) germplasm using microsatellite DNA, with special emphasis on the Fiji Dwarf cultivar. Theor. Appl. Genet. 106:715-726.

Minch, E., Ruitz-Linares, A., Goldstein, D., Feldman, M., Kidd, J.R. and Cavalli-Sforza. 1997. Microsat 1.5: a computer program for calculating various statistics on microsatellite allele data. http://hpgl.stanford.edu/projects/microsat/.

Nei, M. (1972). Genetic distance between populations. Am. Nat., 106:283-292.

Nicese, F.P., Hormaza, J.I. and McGranahan, G.H. 1998. Molecular characterization and genetic relatedness among walnut (*Juglans regia* L.) genotypes based on RAPD markers. Euphytica 101:199-206.

Portis, E., Acquadro, A., Comino, C. and Lanteri, S. 2004. Effect of farmers' seed selection on genetic variation of a landrace population of pepper (*Caspicum annum* L.), grown in North-West Italy. Genet. Resour. Crop Evol. 51:581-590.

Potter, D., Fangyou, G., Aiello, G., Leslie, G. and McGranahan, G. 2002. Intersimple sequences repeat markers for fingerprinting and determining genetic relationships of walnut (*Juglans regia*) cultivars. J. Amer.Soc. Hort. Sci. 127 (1):75-81.

Warburton, M. and Hoisington, D. 2001. Applications of molecular marker techniques to the use of international germplasm collections. p.15-27. In: Henry R.J. (eds.), Plant Genotyping: the DNA fingerprinting of Plants.

Woeste, K., Burns, R., Rhodes, O. and Michler, C. 2002. Thirty polymorphic nuclear microsatellite loci from black walnut. J. Hered. 93:58-60.

Yamamoto, T., Kimura, T., Sawamura, Y., Manabe, T., Kotobuki, K., Hayashi, T., Ban, Y. and Matsuta, N. 2002. Simple sequence repeats for genetic analysis in pear. Euphytica 124:129-137.

Zhao, X. and Kochert, G. 1992. Characterization and genetic mapping of a short, highly repeated, interspesed DNA sequence from rice (*Oryza sativa* L.) Mol. Gen. Genet. 231:353-359.

The use of trade names is for the information and convenience of the reader and does not imply official endorsement or approval by the United States Department of Agriculture or the Forest Service of any product to the exclusion of others that may be suitable.

Tables

Table 1. Name, origin, and code of plants sampled.

Name	Origin	Code
'Bleggiana'	Italian	B
'Franquette'	French	F
'Hartley'	USA	H
'Malizia'	Italian	M
'Parisienne'	French	P
'Serr'	USA	SR

Table 2. Properties of the microsatellite loci used to characterize six walnut genotypes.
[1]Size of a specific allele that is present only in one genotypes.

Locus	T.m. (°C)	Size range (bp)	Total alleles	Private alleles[1] (allele size in bp)
WGA4	50	225 – 239	5	225 = F; 229 = B; 231 = H
WGA33	50	193 – 205	6	195 = P; 199 = H; 203 = M 205 = B
WGA80	50	130 – 188	4	132 / 186 = B
WGA147	47	130 – 198	7	130 = F; 194 = H; 196 = SR 198 = M
WGA148	50	132 – 196	5	134 / 194 = H
WGA204	50	132 – 182	4	132 = B; 160 = F
WGA221	47	252 – 296	8	252 / 256 = B; 258 = H 260 / 276 = SR; 264 = P 296 = F
WGA256	50	200 – 202	2	-
WGA275	50	200 – 206	4	204 = SR; 206 = M
Total			45	26
Mean			5.0	

Table 3(a,b,c). The distribution of the putative alleles for nine polymorphic microsatellite loci in walnut.

Table 3a.

	WGA4					WGA33						WGA80			
	Allele (bp)														
Code	225	229	231	233	239	193	195	199	201	203	205	130	132	186	188
B		+									+		+	+	
F	+			+		+			+			+			+
H			+		+	+		+				+			+
M				+	+	+				+		+			+
P				+	+		+		+			+			+
SR				+	+	+			+			+			+
Freq.(%)	9	9	9	36	36	36	9	9	27	9	9	42	8	8	42

Table 3b.

	WGA147							WGA148				WGA204			
	Allele (bp)														
Code	130	132	136	192	194	196	198	132	134	192	194	196	132	160	162	182
B			+	+				+		+			+			
F	+							+						+		
H		+			+				+		+				+	+
M		+					+	+				+			+	+
P			+	+				+		+					+	+
SR		+				+		+				+			+	+
Freq.(%)	9	27	18	18	9	9	9	45	9	18	9	18	10	10	40	40

Table 3c.

	WGA221								WGA256		WGA275			
	Allele (bp)													
Code	252	254	256	258	260	264	276	296	200	202	200	202	204	206
B	+		+						+		+			
F		+						+	+		+	+		
H				+							+	+		
M									+	+				+
P		+				+				+	+	+		
SR					+		+		+	+	+		+	
Freq.(%)	11	22	11	11	11	11	11	11	57	43	45	36	9	9

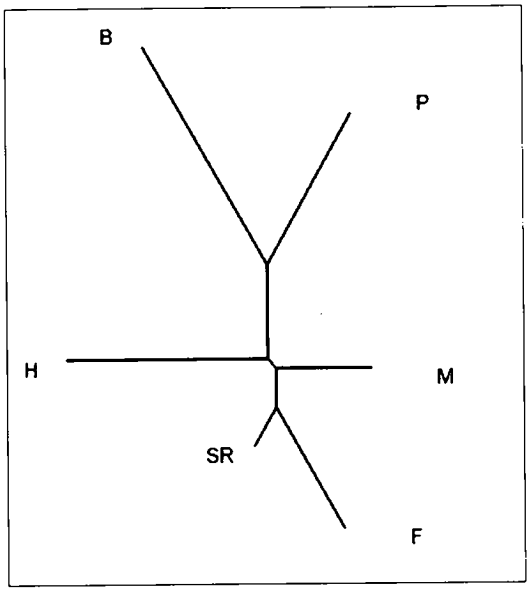

Fig. 1. Unrooted tree representing relationships among six samples obtained using nine polymorphic microsatellites.